AD-A241 232

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

...nated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and ...collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including ...s Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, ...ction Project (0704-0188), Washington, DC 20503.

2. REPORT DATE
August 1991

3. REPORT TYPE AND DATES COVERED
professional paper

4. TITLE AND SUBTITLE
APPLICATIONS OF A CONDITIONAL EVENT ALGEBRA TO DATA FUSION

5. FUNDING NUMBERS
In-house funding

6. AUTHOR(S)
I. R. Goodman

DTIC ELECTE OCT 3 1991 S C D

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Ocean Systems Center
San Diego, CA 92152–5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Ocean Systems Center
San Diego, CA 92152–5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This paper, in the spirit of last year's article, is a contribution toward a unified theory of data fusion as an integral part of a more generic C^3 theory. A computationally feasible and mathematically sound procedure is proposed here for the combination of disparate information, prior to, and compatible with, ordinary conditional probability evaluations. This is based directly upon a new breakthrough concerning the extension of classical probability logic to a full conditional logic.

91-12109

Published in *Technical Proceedings DFS-88, 1988 Tri-Service Data Fusion Symposium*, May 1988.

14. SUBJECT TERMS
probability logic
C^3
data fusion
algebraic logic description pair

Reprints

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

UNCLASSIFIED

21a. NAME OF RESPONSIBLE INDIVIDUAL	21b TELEPHONE *(Include Area Code)*	21c. OFFICE SYMBOL
I. R. Goodman	(619) 553-4014	Code 421

DFS - 88

1988 TRI-SERVICE DATA FUSION SYMPOSIUM

TECHNICAL PROCEEDINGS

VOLUME I

JOHNS HOPKINS UNIVERSITY
APPLIED PHYSICS LABORATORY
LAUREL, MARYLAND

17 - 19 MAY 1988

SPONSORED BY :

JOINT DIRECTORS OF LABORATORIES
DATA FUSION SUB-PANEL
AND
CECOM CENTER FOR SIGNALS WARFARE
NAVAL AIR DEVELOPMENT CENTER
ROME AIR DEVELOPMENT CENTER
NAVAL OCEAN SYSTEMS CENTER

APPLICATIONS OF A CONDITIONAL EVENT ALGEBRA TO DATA FUSION

Dr. I.R. Goodman

Command & Control Department
Code 421
Naval Ocean Systems Center
San Diego, California 92152

Abstract

This paper, in the spirit of last year's article, is a contribution toward a unified theory of data fusion as an integral part of a more general C^3 theory. A computationally feasible and mathematically sound procedure is proposed here for the combination of disparate information, prior to, and compatible with, ordinary conditional probability evaluations. This is based directly upon a new breakthrough concerning the extension of classical probability logic to a full conditional logic.

1. DATA FUSION AND C^3 PROCESSES

1.1. Qualitative aspects.

Previously, in [1] a general approach to data fusion was outlined within the context of C^3 processes.

As a brief review (see also [2],[3]) and a modification of past efforts, the following obtains: The author has been considering C^3/processes from the generic viewpoint of interacting nodes of decision-makers, or complexes of such. These interactions or "signals" may be actual vectors of signals containing voluntary or leaked information from other nodes, friendly or hostile, or they may represent fired weapons, for example. The nodes are relative in size, but whether they represent one or a group of individuals, they possess certain common characteristics. These include a decision structure centering around data fusion which contains detection, hypotheses forming mechanisms, algorithm selections, and responses as an output. The nodes also are represented by corresponding state vectors containing all pertinent descriptors, such as equations of motion and location, number of individuals or supplies and their relative importance, damage levels, threat levels, and estimates and other knowledge of other node states, friendly or adversary.

Schematically, the following simplified situations hold as Figures 1 and 2 show:

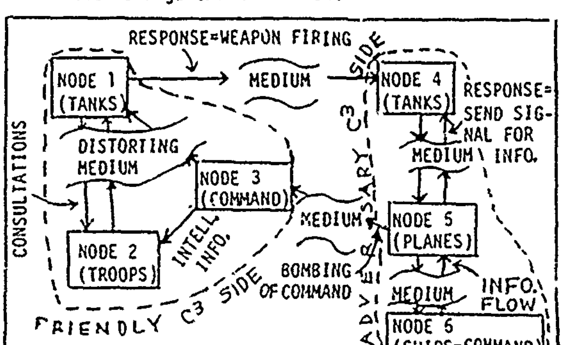

Figure 1. External Dynamics of C^3 Processes: Simplified

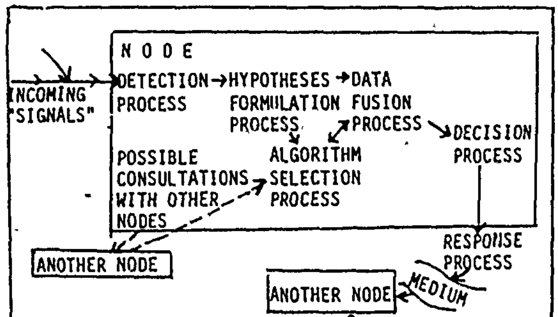

Figure 2. Internal Dynamics of C^3 Processes: Simplified

The basic evolution cycle of a typical node is described in the following figure:

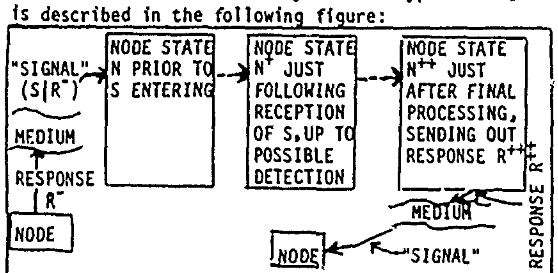

Figure 3. Basic Evolution Cycle of a Node Due to "Signal" Processing and Response.

The components of a typical node state are as in Table 1:

NODE STATE VECTOR	=	NODE STATE PROPER	=	THREAT LEVEL # OF TROOPS # OF WP, I # OF WP, II IMPORTANCE SUPPLY LEVEL EQ. OF MO. DAMAGE LEVEL
		KNOWLEDGE PART		ESTIMATES OF OTHER NODE STATES

Table 1. Components of C^3 Node States.

1.2. Quantitative aspects.

The next step following the qualitative scoping-out of C^3 processes, including data fusion, is the determination of the corresponding quantitative description. In effect, this entails choosing both an appropriate relational syntax and a numerical/semantic evaluation system. Such a pair is called an *algebraic logic description pair* (ALDP). The reader is, most likely,

familiar with the most common ALDP: probability logic (PL) where the appropriate syntax is the structure of a boolean (or more strongly, sigma-) algebra R of events. Here, the usual set or logical/propositional operations hold: unions($\cup$) or disjunctions($\vee$), intersections($\cap$) or conjunctions(&), complements or negations ($()'$), material implications ($(\cdot)'\cup(\cdot\cdot)$ or $(\cdot)'\vee(\cdot\cdot)$, commonly denoted by $\cdot\rightarrow\cdot$), and material double implication or material or logical equivalence ($(\cdot\rightarrow\cdot\cdot)\&(\cdot\cdot\rightarrow\cdot)$, commonly denoted by $\cdot\leftrightarrow\cdot\cdot$). For purposes of simplicity and because the author has propositional logic and its extensions in mind, the common notation used throughout this paper will be the logical/propositional interpretations- but these can all be immediately converted to the set notation counterparts, where required. The corresponding semantic or numerical evaluation for PL is of course simply the choice of a particular (joint) probability measure (either finitely additive or countably additive, if need be) $p:R \rightarrow [0,1]$, the last symbol denoting the unit interval. (See [1] and [4] for background.)

As pointed out in [1] and elsewhere ([4],Chpts 1, 2.2.1), other processes are also involved in establishing an evaluation of a situation such as that entailing C^3 and data fusion processes, besides the choice of ALDP. For example, mental imaging and cognitive processes play an important role,as do natural language formulations, semiotics, and full formal language / formal theory. For simplicity only the latter will be considered in any detail prior to the choice of ALDP. This sequence of knowledge flow in converting any qualitative description into a quantitative one can be summarized in the diagram in Figure 4:

Q U A L I T A T I V E   P H A S E

PROBLEM/SITUATION (C^3 PROCESSES,E.G.) $\longrightarrow$ COGNITION $\longrightarrow$ NATURAL LANGUAGE FORMULATION

Q U A N T I T A T I V E   P H A S E

ALGEBRAIC LOGIC DESCRIPTION PAIR: (SYNTAX,SEMANTICS) $\longleftarrow$ FULL FORMAL LANGUAGE/ FORMAL THEORY $\longleftarrow$ SEMIOTICS/ STRINGS OF SYMBOLS, DICTIONARIES

Figure 4. Knowledge Flow in Describing Situations.

Examples of ALDP's include:

probability logic(PL):(boolean algebra,prob. measure)
fuzzy logic(FL):(browerian lattice, possibility meas.)
Dempster-Shafer logic(DSL):(boolean algebra,belief ms.)
classical logic (CL): (boolean algebra,0-1 valued ms.)
(See [1] or [4] again for further details.)

In the case of C^3 processes, an improved full formal language description has been developed for the dynamic evolution of a typical C^3 node state vector [5], replacing previous efforts in [1]-[3]. The next section describes this.

1.3 Full formal language description/ theory for C^3 node evolutions.

In brief, the full formal language description is summarized in the following tables:

EQUALITY SYMBOL: =
CONSTANTS: Ω, $\emptyset$
DUMMY VARIABLES: α,β,γ
SPECIFIC VARIABLES: N,R,S; T IMPLICIT IN N
OPERATORS: $()^+, ()^-, ()_0, (\,|\,)$, &, $\vee$, DOM, ϵ

GENERAL AXIOMS: FOR ALL α,β,γ, AND FOR $*$ = &,$\vee$:
RING STRUCTURE FOR &,$\vee$:

$$\alpha*\beta = \beta*\alpha\ ,\quad \alpha*(\beta*\gamma) = (\alpha*\beta)*\gamma\ ,$$
$$\alpha\&\emptyset = \emptyset,\ \alpha\&\Omega = \alpha = \alpha\vee\emptyset,\ \alpha\vee\Omega = \Omega,$$
$$\alpha\&(\beta\vee\gamma) = (\alpha\&\beta)\vee(\alpha\&\gamma).$$

Table 2a. Formal Language Description of a C^3 Node Evolution: Part 1.

GENERAL AXIOMS: FOR ALL α,β,γ, AND FOR $*$ = &,$\vee$:
IMPLICATIVE/CONDITIONAL STRUCTURE FOR &,$\vee$:

$$(\alpha|\Omega) = \alpha\ ,\ (\alpha|\beta) = (\alpha\&\beta|\beta)\ ,$$
$$(\alpha\&\beta|\gamma) = (\alpha|\beta\&\gamma)\&(\beta|\gamma),$$
$$(\alpha*\beta|\gamma) = (\alpha|\gamma)*(\beta|\gamma)\ ,$$
$$\bigvee_{\alpha\epsilon DOM(\alpha)} \alpha = \Omega\ .$$

SUFFICIENCY AXIOMS: FOR $\tilde{N} \overset{d}{=} (N\&N^{--}\&N^{==}\&\cdots\&N_0)$:

$$(R^{++}|N^+\&\tilde{N}) = (R^{++}|N^+)$$
$$(N^{++}|R^{++}\&\ N^+\&\ \tilde{N}) = (N^{++}|R^{++}\&\ N^+)$$
$$(N^+|S\ \&\ R^-\&\ \tilde{N}) = (N^+|S\ \&\ N)$$
$$(S|R^-\&\ \tilde{N}) = (S|R^-)$$

Table 2b. Formal Language Description of a Node Evolution: Part 2.

The symbols in Table 2 can be interpreted as follows,as given in Table 3:

N = NODE STATE VECTOR , T = NODE STATE STRUCTURE
R = RESPONSE VECTOR , S = "SIGNAL" VECTOR
$(\ )^+$ = POSITIVE TIME SHIFT TO NEW PHASE
$(\ )^-$ = NEGATIVE TIME SHIFT TO OLD PHASE
$(\ )_0$ = INITIALIZATION OF STATE (TIME-WISE)
$(\ |\ )$ = IMPLICATION OR CONDITIONING
& = AND , $\vee$ = OR , $(\ )'$ = NOT (EXPLAINED EARLIER)
DOM = DOMAIN OF POSSIBLE VALUES
ϵ = SET MEMBERSHIP RELATION AS USED BEFORE

Table 3. Interpretations of the Formal Language for C^3 Node Evolution.

Note that N and T above can be partitioned into subvectors as e.g.:

$$N = (\#WP_1,\#WP_2,\#WP_3,\#TROOP,EQMO,\widehat{INFO}) \qquad (1.1)$$
$$T = (DET,ALG,HYP,FUS,CONS,DEC) \qquad (1.2)$$

where DET=detection (or not), ALG=algorithm selection, HYP=hypotheses formulation, FUS=data fusion,DEC=decision, etc. One can add the constraint (1.3) to the axioms in Table 2b:

$$(R^{++}|T^+\&\ N^+) = (R^{++}|DEC^+\&\ N^+). \qquad (1.3)$$

Using similar sufficiency assumptions, implicative chaining in Table 2b shows that

$$\begin{aligned}(T^+|N^+) = {} & (DEC^+|CONS^+\&FUS^+\&HYP^+\&ALG^+\&DET^+\&N^+)\\ & \&(CONS^+|FUS^+\&HYP^+\&ALG^+\&DET^+\&N^+)\\ & \&(FUS^+|HYP^+\&ALG^+\&DET^+\&N^+)\\ & \&(HYP^+|ALG^+\&DET^+\&N^+)\\ & \&(ALG^+|DET^+\&N^+)\&(DET^+|N^+)\ . \end{aligned} \qquad (1.4)$$

Finally, applying the usual deduction procedure to the axioms given in Table 2, yields the following theorem (1.1)describing the dynamic evolution of a node state in formal language terms:

<u>Theorem</u> 1.1 (See [5].)

Under the assumptions in Table 2:

$$(N^{++}|N) = \bigvee_{\begin{pmatrix} R^{++}\epsilon\ DOM(R^{++}) \\ N^+\ \epsilon\ DOM(N^+) \end{pmatrix}} ((N^{++}|R^{++}\&N^+)\&(R^{++}|N^+)\&(N^+|N)) \qquad (1.5)$$

where

$$(N^+|N) = \bigvee_{R^-\epsilon DOM(R^-)} ((N^+|R^-\&N)\&(R^-|N)), \qquad (1.6)$$
$$(R^{++}|N^+) = \bigvee_{T^+\epsilon DOM(T^+)} ((R^{++}|T^+\&N^+)\&(T^+|N^+)), \qquad (1.7)$$
$$(N^+|R^-\&N) = \bigvee_{S\epsilon DOM(S)} ((N^+|S\&N)\&(S|R^-))\ , \qquad (1.8)$$

and

$$N^{++} = \bigvee_{N\epsilon DOM(N)} ((N^{++}|N)\&\ N)\ . \qquad (1.9)$$

Thus, using the interpretation in Table 3, compatible with Figure 3,

$$(R^{++}|T^{+}\&N^{+}) = \text{response following processing,} \quad (1.10)$$

$$(N^{++}|R^{++}\&N^{+}) = \text{new node state due to its sending out response ,} \quad (1.11)$$

$$(T^{+}|N^{+}) = \text{processing data,} \quad (1.12)$$

$$(N^{++}|N) = \text{full cycle of node change due to "signals" received, over all possible processing, and responses.} \quad (1.13)$$

etc.

Thus, if PL were chosen as the ALDP, assuming only stochastic relations are involved in C^3 variables, then Theorem1.1 reduces to the more familiar form

$$p(N^{++}|N) = \int_{\begin{pmatrix}\text{over all}\\ R^{++}\epsilon DOM(R^{++}),\\ N^{+}\epsilon DOM(N^{+})\end{pmatrix}} p(N^{++}|R^{++},N^{+})\cdot p(R^{++}|N^{+})\cdot p(N^{+}|N)dR^{++}dN^{+}. \quad (1.14)$$

Or, if FL were chosen as the ALDP, assuming only fuzzy relations are involved in C^3 variables, then Theorem1.1 becomes under semantic evaluation

$$poss(N^{++}|N) = \max_{\begin{pmatrix}\text{over all}\\ R^{++}\epsilon DOM(R^{++}),\\ N^{+}\epsilon DOM(N^{+})\end{pmatrix}}\left(\min(poss(N^{++}|R^{++},N^{+}), poss(R^{++}|N^{+}), poss(N^{+}|N))\right), \quad (1.15)$$

One could also choose combinations of PL and FL or other ALDP's in the evaluation aspect. (Again, see [4].)

In turn, utilizing the outputs in Theorem 1.1, together with the choice of ALDP, it is clear that the evolution of node states depend on the determination of the relative primitive relations given as sufficiency axioms in Table 2b. Calling each possible combination of such relations for each C^3 side j, friendly (j=1) or adversary (j=2), $PRIM_j$, one can establish a loss $L(PRIM_1, PRIM_2)$ based directly on Theorem 1.1, thereby establishing a zero sum two"person"C^3 decision game and then procede to analyze the game for values, least favorable strategies, bayes decisions, minimax strategies, etc. (See [5] for further details, where a multidimensional gaussian linear conditional structure is imposed upon the relations yielding feasible computational forms.)

Throughout all of the above calculations, data fusion plays a central role (see again (1.4) and (1.7) where the specific quantitative relations depending upon data fusion are shown). In the next section, motivation is developed for a systematic approach to the fusion proper aspect for disparate information arriving in conditional form.

2. NEED FOR A CONDITIONAL EVENT ALGEBRA IN EVALUATING DATA FUSION

2.1 Introduction.

Noting that many of the relations in Table 2b and subsequent equations are in conditioned form, consider in particular the basic data fusion factor in (1.4)

$$Q \overset{d}{=} (FUS^{+}|HYP^{+}\& ALG^{+} \& DET^{+} \& N^{+}), \quad (2.1)$$

noting that the incoming "signal" S is present through the change of N to N^{+} appearing in the antecedent. Fix throughout the discussion, in the antecedent for Q, an arbitrary combination of possible domain values for HYP^{+}, ALG^{+}, DET^{+}, and N^{+} and hence S, up to some variability.

Consider then the two following simplified examples:

$$Q = (a|b) \vee (a|c) , \quad (2.2)$$

where

a = ship target possible position area updated,
b = track history 1 is the assumed assignment (possibly in error) to the target,
c = track history 2 is the assumed assignment (possibly in error) to the target .

Or, perhaps, Q represents intelligence information to be fused by evaluating the truth of the expression

$$Q = ((a|b) \& (c|d)) \vee (e|f), \quad (2.3)$$

where now

a = a(x) = enemy will move up about x troops tomorrow; x=0,50,100,150.
b = b(y) = it will y tomorrow; y=be clear, snow,rain.
c = c(z) = enemy will use Pass z to approach us; z=I,II,III,IV.
d = d(r,s)= morale of enemy node 17 is at level r and number of their troops left is s ; r=very low, low, medium,high, very high, s=0,100,200,300.
e = e(w) = enemy will w tomorrow; w=surrender,not surrender.
f = f(q) = enemy overall damage level is q ; q=0,1,2,..,10.

Of course, if the antecedents in (2.2) or (2.3) were all the same, then no real problem would arise, since for example it is readily justified that for any choice of ALDP - certainly for PL ! - that

$$Q = ((a|d) \& (c|d)) \vee (e|d) = (((a\&c)\vee e)|d) , \quad (2.4)$$

even though normally one does not talk about such *measure-free* entities (up to now). Indeed, since the goal is the evaluation of Q, for PL, choosing a probability measure p over all the relevant events, one would usually evaluate Q as simply

$$p(Q) = p(((a|d) \& (c|d)) \vee (e|d)) = p(((a\&c)\vee e)|d) = p(((a\&c)\vee e))\&d)/p(d), \quad (2.5)$$

etc., assuming p(d)>0.

But the point of the above examples in (2.2) and (2.3) is that the antecedents in the conditional forms are not in general identical! What to do?

Contrary to popular belief [author's note: this author and his colleague Prof. H.T. Nguyen, Math. Dept. New Mexico State University, Las Cruces, have undertaken an extensive informal survey of the probability community- both applied and theoretical-resulting in the following conclusions - see also [7]]: there is no systematic and mathematically sound procedure for computing p(Q) (or Q, for that matter) in either (2.2) or (2.3), or any similar problem!

Indeed, there do exist "folk" remedies to this situation which roughly speaking reduce to two approaches as presented in the next sections.

2.2 Approach A: Identification of conditioning with material implication.

In this approach, one interprets (a|b) as $b \to a$, i.e., for any a,b,events,

$$(a|b) = b\to a \overset{d}{=} b'\vee a = b'\vee(a\&b) \quad (2.6)$$

is assumed to be valid so that by the principle of substitution relative to equality, for any (suitable) probability measure p, assuming the identity for p(b)>0

$$p((a|b)) = p(a|b) \overset{d}{=} p(a\&b)/p(b), \quad (2.7)$$

one has immediately from (2.6)

$$p(a|b) = p(b\to a) = p(b'\vee(a\&b)) = 1-p(b)+p(a\&b), \quad (2.8)$$

etc.

Thus using this approach, (2.2) becomes:

$$p(Q) = p((b\to a) \vee (c\to a)) = p(b'\vee a\vee c'\vee a) = p((b\&c)'\vee a) = p((b\&c)\to a) = p(a|b\&c). \quad (2.9)$$

However, there is just one flaw in the above reasoning: One cannot make the identification in general in the left hand side of (2.6). Indeed, it is rather easy to show (yet surprisingly few are aware of the inequality below- see e.g. the discussion in [7], sect. 1.8):

$$p(b\to a) = p(a|b) + p(a'|b)\cdot p(b') \geq p(a|b) \geq p(a\&b) , \quad (2.10)$$

where in general strict inequality holds for the last two ≥. Furthermore, (2.6) is not even a good approximation, since it is readily verified for b with p(b) small that one can choose a and b such that $p(b \Rightarrow a)$ is close to one, while $p(a|b)$ is close to zero. This unfortunate situation is a special case of the "Stalnaker Thesis" problem and is considered in detail in [6] and [7], section 1. In fact the following apparently not-well-known result has been shown:

<u>Theorem</u> 2.1. (P. Calabrese [8], section 1.2)

Let R be a boolean algebra of events with the usual operations discussed previously. Let p be any non-degenerate probability measure over R, i.e. $p:R \to [0,1]$. Then, there is <u>no</u> binary boolean function $f:R^2 \to R$ such that for all $a,b \in R$, with $p(b)>0$,

$$p(a|b) = p(f(a,b)). \qquad (2.11)$$

∎

The result is further extended in one direction by the following result:

<u>Theorem</u> 2.2. (Goodman & Nguyen [7], Theorem 2.7)

Let R be any <u>finite</u> boolean algebra of events. Then, there is no <u>binary function</u> f of any kind, $f:R^2 \to R$, such that for all $a,b \in R$, with $p(b)>0$, (2.11) holds.

∎

However, by allowing infinite boolean algebras one can force a form of conditioning to be back in the original boolean algebra (but not without complication) as Copeland showed. (For a critique of Copeland's "implicative" boolean algebras, see [7], section 1.8.)

2.3 <u>Approach B: Identification of conditional events as marginal ones with common joint antecedents.</u>

In this approach, one attempts to obtain a common joint antecedent event for all of the conditional events appearing and then identify each as a marginal one having a common joint antecedent. In turn, one procedes to evaluate for a suitably chosen probability measure analogous to that in (2.5). This is best shown through an application to (2.3):

First make the identifications

$$\left.\begin{array}{l}(a|b) \text{ with } (a\times d\times f|G)\\ (c|d) \text{ with } (b\times c\times f|G)\\ (e|f) \text{ with } (b\times d\times e|G)\end{array}\right\} ; G \overset{d}{=} b\times d\times f. \qquad (2.12)$$

Then (2.3) becomes

$$Q = ((a\times d\times f|G) \;\&\; (b\times c\times f|G)) \vee (b\times d\times e|G)$$
$$= (((a\&b \times c\&d \times f) \vee (b\times d\times e))|G) , \qquad (2.13)$$

resulting in the evaluation

$$p(Q) = p_0((a\&b \times c\&d \times f) \vee (b\times d\times e) \mid G) , \qquad (2.14)$$

which can be further reduced using the definition of conditional probability and the standard calculus of operations for PL.

However, the main drawback to this approach is that initial probability measure $p:R \to [0,1]$ must be replaced by some joint probability measure p_0 over the boolean (or sigma-) algebra spanned by R^3, where $p(\cdot|b), p(\cdot\cdot|d), p(\cdots|f):R \to [0,1]$ are conditional probability measures whose joint measure is p_0, so that

$$p(a|b)=p_0(a\times d\times f|G), p(c|d)=p_0(b\times c\times f|G), p(e|f)=p_0(b\times d\times e|f) \qquad (2.15)$$

But what choice of p_0 to make? Should it be based on maximal entropy considerations, etc.? Long calculations can also result from the cartesian product forms. (For further discussion, see [7], sections 1.1, 1.5, and 10.4. However, for a tie-in with the approach presented in this paper, Theorem 4.2 given in section 4 is of use.)

2.4 <u>The basic problem in representing conditional events with distinct antecedents.</u>

As shown in section 2.3 two of the most common approaches to the handling of conditional events do not lead to satisfactory results, from either a mathematical viewpoint, as in Approach A, or an unambiguous computationally efficient viewpoint, as in Approach B.

Thus one is lead to question whether any remedy exists for this situation: Can a calculus of measure-free conditional events be developed which is both compatible with ordinary conditional probability evaluations and is also unambiguous and feasible to implement, as well as being based upon sound, non-ad hoc mathematical principles? Certainly, all evaluation of fused data, and hence evaluations for the overall C^3 problem, must depend, in effect, on the outcome of this question.

It is the contention here that the answer to the above question is definitely in the affirmative. In this paper in section 3 an outline of a theory for developing such a calculus of operations and related issues is presented. This is carried out for not only the purpose of keeping this paper as self-contained as possible, but also because of the desire to disseminate these novel and universally applicable results to as wide an audience as possible, within a short time period. For earlier efforts in this direction, see [6]. In [7](in the process of being submitted for publication) the full theory, with all proofs, is exhibited.

One consequence of the calculus of conditional events is that the evaluations for Q in (2.2) and (2.3) become rather simple. Thus, it will be shown that (2.2) yields

$$Q = ((a\&(b\vee c)|(a\&(b\vee c))\vee(b\&c)) \qquad (2.16)$$

so that

$$p(Q) = p(a\&(b\vee c))/p((a\&(b\vee c))\vee(b\&c)), \qquad (2.17)$$

etc., differing considerably from that proposed by Approaches A (see (2.9)) or B.

For (2.3), one obtains

$$Q = (\alpha|\beta) \qquad (2.18)$$

resulting in

$$p(Q) = p(\alpha)/p(\beta), \qquad (2.19)$$

where

$$\alpha \overset{d}{=} (a\&b\&c\&d) \vee (e\&f) , \qquad (2.20)$$
$$\beta \overset{d}{=} \alpha \vee (((a'\&b)\vee(c'\&d)\vee(b\&d))\&f), \qquad (2.21)$$

differing again considerably from both Approaches A and B (see (2.13)).

In all of the above computations, one need not construct joint probability measures and consider cartesian products of events, nor is the procedure ad hoc (despite the oversimplified appeal of Approach A - but see section 4(9)).

3. OUTLINE OF A THEORY OF MEASURE-FREE CONDITIONAL EVENTS

3.1 <u>Introduction.</u>

Following the basic motivation for the development of a conditional event calculus for PL in section 2, this section presents an overview of the basic results. The following three questions are addressed:

(i) What meaning can be attached to a typical conditional event (a|b), where a and b are ordinary unconditional events, prior to evaluating through a specific probability measure p to become (for $p(b)>0$)

$$p((a|b)) = p(a|b) \overset{d}{=} p(a\&b)/p(b) \;? \qquad (3.1)$$

(ii) How shall operations- in particular, the usual boolean operations &,v,()' and relations such as ≤ - be extended from unconditional events to conditional ones and what properties do they possess?

(iii) Can such operations and relations as in (ii) be characterized for uniqueness, etc. ?

For a history of previous attempts at developing a theory of measure-free conditioning, see [7], sections 1.8 and 1.9. Among the predecessors of this effort, Schay [9] was among the first to attempt such a task, but used an ad hoc procedure in addressing question (ii), although a somewhat complicated characterization was developed relative to (iii). Later, Calabrese, completely independent of Schay, produced also an algebra of conditional events and operations,

the latter also from an ad hoc approach following certain analogies with material implication [8].

3.2 Development of measure-free conditional events.

From now on, without further explanation, the symbol R refers to an arbitrary fixed boolean algebra with all of the usual operations and relations explained previously (see section 1.2). The partial order $\le$ over R^2 (corresponding to the subset relation $\subseteq$ among sets corresponding to events or propositions as considered here) is defined as usual as

$$a \le b \text{ iff } a = a\&b \text{ iff } b = a v b, \tag{3.2}$$

for any (unconditional) events $a,b \in R$. Note also, $\emptyset$ denotes the null event and Ω the universal (unity)event.

Define (f,S) to be any *candidate class of conditional events extending* R iff f is some function $f:R^2 \to S$ for some space S, such that for all $a,b,c \in R$, and boolean operations $*$ over R: $*_b$, etc. are operations over (assumed) boolean algebra of events

$$S_b \overset{d}{=} f(\cdot,b)(R) = \text{range}(f(\cdot,b)) \tag{3.3}$$

where

$$S = \bigcup_{b \in R} S_b, \tag{3.4}$$

$$f(a,b) *_b f(c,b) = f(a*c,b), \tag{3.5}$$

and

$$f(a,b) = f(a\&b,b), \quad f(a,\Omega) = a. \tag{3.6}$$

For any candidate class (f,S), one can interpret $f(a,b)$ as a conditioned upon b wrt f or symbolically as $(a|b)_f$, where a is the consequent and b the antecedent.

Theorem 3.1. ([7], Theorem 2.1)

(i) $(\text{nat},\tilde{R})$ is a candidate class of conditional events extending R, called the *canonical class of conditional events extending* R, where in (3.4)

$$R_b \overset{d}{=} \{(a|b)_{nat}: a \in R\}, \tag{3.7}$$

where for each $a,b \in R$,

$$(a|b) \overset{d}{=} (a|b)_{nat} \overset{d}{=} (R\&b')v(a\&b) \overset{d}{=} \{(x\&b')v(a\&b): x \in R\} \tag{3.8}$$

is the principal ideal coset ($R\&b'$ being the principal ideal) generated by b' with residue a&b, making R_b a boolean quotient algebra with the usual coset operations $\tilde{*}_b$ corresponding to ordinary boolean operations $*$ over R (thus, $* = \&, v, ()'$, etc.).

$$\tilde{R} \overset{d}{=} \bigcup_{b \in R} R_b = \{(a|b): a,b \in R\} \tag{3.9}$$

and for any $a,b \in R$,

$$\text{nat}(a,b) \overset{d}{=} (a|b). \tag{3.10}$$

(ii) For each $a,b \in R$, noting $(a|b) \subseteq R$, $(a|b)$ is the inverse of the conjunction operation $(\cdot\&b):R \to R$ at a&b, i.e.,

$$(a|b) = \{y: y \in R \text{ and } y\&b = a\&b\}. \tag{3.11}$$

Equivalently, $(a|b)$ is the largest subclass of R satisfying the relation

$$(a|b)\&b = a\&b \tag{3.12}$$

(iii) For any candidate class (f,S) of conditional events extending R, and each $b \in R$,

$$S_b(\&_b,v_b,()'^b) \approx_b R/f^{-1}(\cdot,b) \text{ (boolean quot.alg.)} \supseteq R_b(\tilde{\&}_b,\tilde{v}_b,()^{\tilde{\prime}b}), \tag{3.13}$$

the symbol $\approx_b$ denoting an isomorphism, so that in the above local sense for each b, the canonical class of conditional events is the smallest possible candidate class of conditional events extending R. ∎

Note also that for any $a,b,c,d \in R$,

$$(a|b) = (c|d) \text{ iff } a\&b = c\&d \text{ and } b = d, \tag{3.14}$$

but unlike the classical case where b=d, (a|b) is not necessarily identical nor disjoint from (c|d). (See [7], Theorem 2.11.) Note also the relations

$$R \le \tilde{R} \le P(R), \tag{3.15}$$

where $P(R)$ denotes the power class, or class of all subsets of elements, of R.

Next, call any pair (f,S) (not necessarily a priori a candidate class of conditional events) where $f:R^2 \to S$ is surjective (i.e., onto) a *probability-compatible pair with respect to* R iff, by definition,

$$R \subseteq S = \bigcup_{b \in R} S_b \tag{3.16}$$

where now S_b is some boolean algebra, for each $b \in R$,

and where for each probability measure $p:R \to [0,1]$ and each $b \in R$ with $p(b)>0$, p can be extended to (using the same symbol) $p:S \to [0,1]$ such that for all $a \in R$,

$$p(f(a,b)) = p(a|b). \tag{3.17}$$

Theorem 3.2. ([7], Theorems 2.5, 2.9)

Let (f,S) be a probability-compatible pair with respect R. Then:

(i) (f,S) is also a candidate class of conditional events extending R.

(ii) $(\text{nat},\tilde{R})$ is a probability-compatible pair.

(iii) By suitable restriction of $f:R^2 \to S$ to nontrivial, i.e., non-zero or non-unity probability-valued conditional event pairs, then : f becomes bijective and $\approx_b$ in (3.13) can be replaced by a global isomorphism $\approx$ not dependent upon b, showing that the canonical class of conditional events $(\text{nat},\tilde{R})$ is the universally minimal (wrt subclass inclusion) probability-compatible candidate class of conditional events extending R. ∎

Next, a basic logical justification for choosing the canonical conditional events is given. First, for any $a,b \in R$, define the class of all a-*relative deducts of* b as

$$\{a|b\} \overset{d}{=} \{x: x \in R \text{ and there exists } r \in R \text{ with } r \ge b \text{ and } x\&r = a\&r\} \tag{3.18}$$

(Calabrese [8], motivated by classical logic deduction, proposed this definition previously for conditioning a by b.)

Theorem 3.3, ([7], Theorem 1.3)

For all $a,b \in R$,

$$\{a|b\} = (a|b), \tag{3.19}$$

i.e., $\tilde{R}$ coincides with the class of all relative deducts! ∎

3.3 Calculus of operations on conditional events.

With the role of the canonical class of conditional events firmly established in section 3.2, consider now the choice of operations upon them extending the usual boolean operations acting upon unconditional events. Throughout all of real analysis and topology, a universal way of extending a given "point"-valued function to a "set"-valued one is simply through the *natural extension* or, equivalently, called the *component-wise class or image extension*. In particular, let $g:R^2 \to R$ be any binary operation (boolean or otherwise) upon R (unary and more generally, n-ary operations can be treated similarly). Then the natural extension (also denoted by the same symbol for g) from g over R^2 to $g:P(R)^2 \to P(R)$, i.e., over $P(R)^2$, is determined by,

$$g(A,B) \overset{d}{=} \{g(a,b): a \in A, b \in B\}, \tag{3.20}$$

for any $A,B \subseteq R$, i.e., $A,B \in P(R)$.

Thus, recalling the comments in Theorem 3.1(ii) and (3.15), it is basic to inquire what forms the natural extensions take for the ordinary boolean operations relative to R when restricted to $\tilde{R}$. Note that for the binary operations $\&,v:R^2 \to R$, by their commutativity and associativity: they extend recursively

and unambiguously to $\&,\vee:\mathcal{R}^n \to \mathcal{R}$, where, by convention, for $n=1$, $\& = \vee =$ identity function. (The unary negation or complement operator $()':\mathcal{R} \to \mathcal{R}$ remains as is.) Denote the natural extensions of $\&,\vee,()'$ by the same symbols. Thus, here the behavior of the natural extensions of $\&,\vee:\tilde{\mathcal{R}}^n \to \tilde{\mathcal{R}}$ to the restrictions $\&,\vee:\tilde{\mathcal{R}}^n \to \mathcal{P}(\mathcal{R})$ and that of $()':\mathcal{R} \to \mathcal{R}$ to the restriction $()':\tilde{\mathcal{R}} \to \mathcal{P}(\mathcal{R})$ are sought.

<u>Theorem</u> 3.4. ([7], Corollaries 3.3,3.4)

For any positive integer n and any $a_j, b_j \in \mathcal{R}$, $j=1,\ldots,n$:

(i) The common antecedent case, noting $\mathcal{R}_b \subseteq \mathcal{R}$.

$$(a|b)' = (a|b)^{\prime_b} = (a'|b) = (a'\&b|b), \qquad (3.21)$$

$$\overset{n}{\underset{j=1}{*}}(a_j|b) = \overset{n}{\underset{j=1}{*_b}}(a_j|b) = (\overset{n}{\underset{j=1}{*}} a_j|b) = (\overset{n}{\underset{j=1}{*}} a_j\&b|b), \qquad (3.22)$$

for $* = \&,\vee$.

Thus, relative to any given boolean quotient algebra of the form $\mathcal{R}_b$, all coset operations and corresponding natural extensions of the original operations relative to $\mathcal{R}$ coincide. It follows that if $g:\mathcal{R}^n \to \mathcal{R}$ is any (compound) boolean operation, then

$$g((a_1|b_1),\ldots,(a_n|b_n)) = \bar{g}_b((a_1|b_1),\ldots,(a_n|b_n))$$
$$= (g(a_1,\ldots,a_n)|b). \qquad (3.23)$$

(ii) Distinct antecedents, in general, case.

$$\overset{n}{\underset{j=1}{\&}}(a_j|b_j) = (\overset{n}{\underset{j=1}{\&}} a_j|r) = (\overset{n}{\underset{j=1}{\&}}(a_j\&b_j)|r), \qquad (3.24)$$

where

$$r \overset{d}{=} \overset{n}{\underset{j=1}{\&}} b_j \vee \overset{n}{\underset{j=1}{\vee}}(a_j'\&b_j) = \overset{n}{\underset{j=1}{\&}}(a_j\&b_j) \vee \overset{n}{\underset{j=1}{\vee}}(a_j'\&b_j) \qquad (3.25)$$

and

$$\overset{n}{\underset{j=1}{\vee}}(a_j|b_j) = (\overset{n}{\underset{j=1}{\vee}} a_j|q) = (\overset{n}{\underset{j=1}{\vee}}(a_j\&b_j)|q), \qquad (3.26)$$

where

$$q \overset{d}{=} \overset{n}{\underset{j=1}{\vee}}(a_j\&b_j) \vee \overset{n}{\underset{j=1}{\&}} b_j = \overset{n}{\underset{j=1}{\vee}}(a_j\&b_j) \vee \overset{n}{\underset{j=1}{\&}}(a_j'\&b_j). \qquad (3.27)$$

∎

Combining (3.24) and (3.26), leads to the following corollary:

<u>Corollary</u> 3.1 (New result.)

Noting that any (compound) boolean function of multiple arguments can always be put into an equivalent form consisting of a disjunction of conjunctions, let $a_{ij}, b_{ij} \in \mathcal{R}$, $i=1,\ldots,m$, $j=1,\ldots,n$, where some of these events may be $\emptyset$ or Ω. Then

$$\overset{m}{\underset{i=1}{\vee}}\ \overset{n}{\underset{j=1}{\&}}(a_{ij}|b_{ij}) = (\alpha \mid \beta), \qquad (3.28)$$

where

$$\alpha \overset{d}{=} \overset{m}{\underset{i=1}{\vee}}\ \overset{n}{\underset{j=1}{\&}}(a_{ij}\&b_{ij}),\quad \beta \overset{d}{=} \alpha + \gamma, \qquad (3.29)$$

$$\gamma \overset{d}{=} \overset{n}{\underset{j=1}{\&}}\ \overset{m}{\underset{i=1}{\vee}}(a_{ij}'\&b_{ij}), \qquad (3.30)$$

noting that α and γ are disjoint, i.e.,

$$\alpha \,\&\, \gamma = \emptyset, \qquad (3.31)$$

so that for any probability measure $p:\mathcal{R} \to [0,1]$,

$$p(\overset{m}{\underset{i=1}{\vee}}\ \overset{n}{\underset{j=1}{\&}}(a_{ij}|b_{ij})) = p(\alpha|\beta) = p(\alpha)/p(\beta) = 1/(1+(p(\gamma)/p(\alpha))). \qquad (3.32)$$

∎

<u>Theorem</u> 3.6. Some special cases ([7], section 4.2).

For any $a,b,c,d,a_i \in \mathcal{R}$, $i=1,\ldots,n$:

$$(a|\Omega)=a,\ (a|\emptyset)=\mathcal{R},\ (\emptyset|b)=(b'|b)=\mathcal{R}\&b'=\{x\&b': x\in\mathcal{R}\}, \qquad (3.33)$$

$$(\Omega|b)=(b|b)=\mathcal{R}\vee b'=\{x\vee b': x\in\mathcal{R}\}, \qquad (3.34)$$

$$(a|b)\&(c|d)=(a\&b\&c\&d|(a'\&b)\vee(c'\&d)\vee(b\&d))$$
$$=(a\&b\&c\&d|(a'\&b)\vee(c'\&d)\vee(a\&b\&c\&d)), \qquad (3.35)$$

$$(a|b)\vee(c|d)=((a\&b)\vee(c\&d)|(a\&b)\vee(c\&d)\vee(b\&d))$$
$$=((a\&b)\vee(c\&d)|(a\&b)\vee(c\&d)\vee(a'\&b\&c'\&d)), \qquad (3.36)$$

$$(a|b)=(a\&b|b)=(b\to a|b)=(b\to a)\&(b|b)=((b'|a')\vee a)\&(b|b), \qquad (3.37)$$

$$(a|b)\&b=a\&b,\quad (a|b)\vee b=(b|b), \qquad (3.38)$$

$$(a|b)\vee(b|a)=(a\&b|a\&b),\quad (a|b)\&(b|a)=(a\&b|a\vee b), \qquad (3.39)$$

$$(a|b)\vee(c|b')=((a\&b)\vee(c\&b')|(a\&b)\vee(c\&b')), \qquad (3.40)$$

$$(a|b)\&(c|b')=(\emptyset|(a'\&b)\vee(c'\&b')), \qquad (3.41)$$

$$(c|d)\to(a|b)=(c\to a|(c'\&d)\vee(a\&b)\vee(b\&d)), \qquad (3.42)$$

$$(c|d)\leftrightarrow(a|b)=(c\leftrightarrow a|b\&d), \qquad (3.43)$$

chaining:

$$(a\&b|c)=(a|b\&c)\&(b|c), \qquad (3.44)$$

$$(a\&b\&c|d)=(a|b\&c\&d)\&(b|c\&d)\&(c|d), \qquad (3.45)$$

Bayes' Thm:

$$(a_i|b)=((b|a_i)\&a_i|\overset{n}{\underset{j=1}{\vee}}((b|a_j)\&a_j)), \qquad (3.46)$$

provided that

$$b \le \overset{n}{\underset{j=1}{\vee}} a_j, \qquad (3.47)$$

where above, material implication and material double implication, as extended from $\mathcal{R}$ to $\tilde{\mathcal{R}}$, are defined formally the same as in the unconditional case:

$$(c|d)\to(a|b) \overset{d}{=} (c|d)'\vee(a|b), \qquad (3.48)$$

$$(c|d)\leftrightarrow(a|b) \overset{d}{=} ((c|d)\to(a|b))\&((a|b)\to(c|d)). \qquad (3.49)$$

∎

Recalling the partial order $\le$ for $\mathcal{R}$ (see (3.2)), define and extend this relation to $\tilde{\mathcal{R}}$ by letting, for all $a,b,c,d \in \mathcal{R}$,

$$(a|b) \le (c|d)\ \underline{\text{iff}}\ (a|b) = (a|b)\&(c|d). \qquad (3.50)$$

<u>Theorem</u> 3.7. ([7], Theorems 4.1,4.7)

For all $a,b,c,d \in \mathcal{R}$:

(i) Characterization of $\le$ over $\tilde{\mathcal{R}}^2$:

$$(a|b) \le (c|d)\ \underline{\text{iff}}\ (c|d) = (a|b)\vee(c|d)$$
$$\underline{\text{iff}}\ (a\&b)\le(c\&d) \text{ and } (b\to a)\le(d\to c)$$
$$\underline{\text{iff}}\ (a\&b)\le(c\&d) \text{ and } (c'\&d)\le(a'\&b)$$
$$\underline{\text{iff}}\ (c|d)' \le (a|b)'. \qquad (3.51)$$

(ii) $\le$ over $\tilde{\mathcal{R}}^2$ is not only a partial order (reflexive, transitive, anti-symmetric), but also a meet (&) and join($\vee$) lattice with all of the usual operation-preserving properties. Note the relation, compatible with (2.10) and (3.8), showing that $b\to a$ and $a\&b$ are the largest and smallest elements in $(a|b)$, respectively, relative to partial order $\le$:

$$a\&b \le (a|b) \le b\to a. \qquad (3.52)$$

(iii) $\tilde{\mathcal{R}}(\&,\vee,()';\le)$ is a distributive lattice which is also an algebraic semi-ring with zero element $\emptyset$ and unity element Ω, and is, further, idempotent, absorbing, and demorgan, among other properties.

∎

3.4 <u>Justifications for choice of extensions of operations from $\mathcal{R}$ to $\tilde{\mathcal{R}}$.</u>

In addition to the large number of desirable properties for the naturally extended boolean operations upon $\tilde{\mathcal{R}}$ given in section 3.3, characterizations can also be established.

Recalling the maximality property of $b\to a$ with respect to $(a|b)$ ((3.52)), define a corresponding mapping $\phi:\tilde{\mathcal{R}} \to \mathcal{R}$, where for any $a,b \in \mathcal{R}$,

$$\phi((a|b)) \overset{d}{=} b\to a. \qquad (3.53)$$

<u>Theorem</u> 3.8. ([7], Theorem 10.1, Remark 10.1)

(i) ϕ is a surjective $\&,\vee$-homomorphism with respect to the natural extensions of $\&,\vee$ from $\mathcal{R}$ to $\tilde{\mathcal{R}}$.

(ii) Let $\tilde{\&},\tilde{\vee}:\tilde{\mathcal{R}}^2 \to \tilde{\mathcal{R}}$ be any possible extensions of $\&,\vee:\mathcal{R}^2 \to \mathcal{R}$, respectively, such that there exist

functions $\psi_1,\psi_2:R^4 \to R$, where for all $a,b,c,d \in R$,

$$a\&b \vee c\&d \le \psi_2(a,b,c,d),\quad a\&b\&c\&d \le \psi_1(a,b,c,d) \tag{3.54}$$

and

$$(a|b)\bar{\&}(c|d) = (a\&b\&c\&d|\psi_1(a,b,c,d)), \tag{3.55}$$

$$(a|b)\mathrm{v}(c|d) = ((a\&b)\mathrm{v}(c\&d)|\psi_2(a,b,c,d)), \tag{3.56}$$

noting the essential necessity of consequences matching the corresponding ones for the natural extensions.

Then it follows that relative to $\tilde{R}$,

$$\breve{\&} = \&\ ,\quad \breve{\mathrm{v}} = \mathrm{v}\ . \tag{3.57}$$

(iii) Compatible with Theorems 2.2, 3.1, 3.2, there is no full (i.e. &,v,()')homomorphism $\rho:\tilde{R} \to R$ with respect to the natural extensions of &,v,()' from R to $\tilde{R}$. ∎

Theorem 3.9. Partial converse of Theorem 3.8. ([7],Theorem 10.2, Corollary 10.1, Remark 10.2)

Let $\breve{\&},\breve{\mathrm{v}}:\tilde{R}^2 \to \tilde{R}$ be any possible extensions of the corresponding coset operations $\breve{\&}_b,\breve{\mathrm{v}}_b:R_b^2 \to R_b$, simultaneously and consistently for all $b \in R$. (Hence, necessarily, $\breve{\&},\breve{\mathrm{v}}$ extend &,v relative to R.)

Suppose also that $\breve{\&},\breve{\mathrm{v}}$ obey not only closure, but are also commutative and associative over $\tilde{R}$ with $\emptyset$ and Ω playing the usual roles of zero and unity elements, respectively. Suppose also there exists $\psi_1,\psi_2:R^4 \to R$ such that for all $a,b,c,d \in R$,

$$(a|b)\breve{\&}(c|d) = (a\&b\&c\&d|\psi_1(a,b,c,d))\ , \tag{3.58}$$

$$(a|b)\breve{\mathrm{v}}(c|d) = ((a\&b)\mathrm{v}(c\&d)|\psi_2(a,b,c,d)), \tag{3.59}$$

$$(a|b)\breve{\&}b = a\&b\ , \tag{3.60}$$

and

$$(\ (a|b)\breve{\&}c \in R\ )\ \text{implies}\ \ c\le b. \tag{3.61}$$

Then

(i) For all $a,b,c,d \in R$,

$$(a'\&b)\mathrm{v}(c'\&d)\mathrm{v}(b\&d) \le \psi_1(a,b,c,d), \tag{3.62}$$

$$(a\&b)\mathrm{v}(c\&d)\mathrm{v}(b\&d) \le \psi_2(a,b,c,d), \tag{3.63}$$

showing that the natural extensions of &,v are maximal, i.e.,

$$(a|b)\breve{\&}(c|d) \le (a|b)\&(c|d), \tag{3.64}$$

$$(a|b)\breve{\mathrm{v}}(c|d) \le (a|b)\mathrm{v}(c|d). \tag{3.65}$$

(ii) Result (i) shows that in contradistinction to the antecedent-only dependent operations of Schay [9] and Calabrese [8], there are no boolean functions ψ_1,ψ_2 as above, but now such that for all $a,b,c,d \in R$, $\psi_j(a,b,c,d)=\psi_j(b,d)$ only, $j=1,2$, such that the corresponding operations $\breve{\&}$ and $\breve{\mathrm{v}}$ satisfy the hypotheses of this theorem. ∎

3.5 Additional properties of $\tilde{R}$.

Finally, brief mention should be made of other pertinent properties of the conditional extension of a given boolean algebra of events.

(i) Stone's Representation Theorem - showing an order-preserving isomorphism always exists between a given boolean algebra R of events or propositions and a corresponding boolean algebra of subsets of some set - can be extended quite readily to $\tilde{R}$ ([7], Theorem 10.3).

(ii) The usual Hilbert-Ackermann axioms involving material implication, relative to any R, when R is replaced by $\tilde{R}$, ordinary substitution, and a modified form of modus ponens used to deduce theorems from previous theorems and axioms, together with the natural extensions of all boolean operations from R to $\tilde{R}$, forms a sound and complete logic. I.e., all theorems are tautologies and vice-versa. Call this *conditional probability logic* (CPL), extending ordinary probability logic (PL). (See [7], Corollary 9.1.) Here, for any $a,b \in R$, $(a|b)$ is a *tautology for* CPL, written $\models_{CPL}(a|b)$, iff, by definition,

$$p(a|b) = 1, \tag{3.66}$$

for all probability measures $p:R \to [0,1]$, with $p(b)>0$.

From [7], sections 9.4, 9.5, 2.5, the following concepts and results hold, for any $a,b,c,d \in R$: (3.67).

(I) $\vdash_{CPL}(a|b)$ iff $b\le a$ iff $b\Rightarrow a = 1$ iff $(a|b)=(b|b)$.

(II) $\{\vdash_{CPL}(a|b)$ and $\vdash_{CPL}(b|a)\}$ iff $a=b$ iff $a\Leftrightarrow b = 1$ iff $(a|b)=(a|a)=(b|b)$. (3.68)

(III) Define $(c|d)$ *tautologically implies* $(a|b)$ iff

$$\models_{CPL}((c|d)\Rightarrow(a|b)). \tag{3.69}$$

Then $\vdash_{CPL}((c|d)\Rightarrow(a|b))$ iff $b\&c\&d \le a$ iff $c\&d \le b\Rightarrow a$. (3.70)

(IV) Define $(c|d)$ *tautologically is equivalent to* $(a|b)$ iff

$$\vdash_{CPL}((c|d)\Leftrightarrow(a|b)). \tag{3.71}$$

Then $\vdash_{CPL}((c|d)\Leftrightarrow(a|b))$ iff $b\&d \le (a\&b\Leftrightarrow c\&d)$. (3.72)

(V) $(c|c)\le(a|b)$ implies $\vdash_{CPL}((c|d)\Rightarrow(a|b))$. (3.73)

(VI) $(c|d)=(a|b)$ implies $\vdash_{CPL}((c|d)\Leftrightarrow(a|b))$. (3.74)

(VII) Sufficient conditions for modus ponens analogue:

$(c|d)\le(a|b)$ and $\vdash_{CPL}(c|d)$ implies $\vdash_{CPL}(a|b)$. (3.75)

(VIII) Characterization of modus ponens analogue:

$(\vdash_{CPL}((c|d)\Rightarrow(a|b))$ and $\vdash_{CPL}(c|d))$ (3.76)

iff $(\vdash_{CPL}(a\&c|b\&d)$ and $\vdash_{CPL}(c|d))$.

(IX) $\vdash_{CPL}((a|b)\Leftrightarrow(c|d))$ implies (3.77)

$(\vdash_{CPL}((a|b)\Rightarrow(c|d))$ and $\vdash_{CPL}((c|d)\Rightarrow(a|b)))$.

(X) Next, define $(c|d)$ *semantically* (or *uniformly in probability*) *implies* $(a|b)$, written $(c|d) \to_{CPL} (a|b)$, iff

$$p(c|d) \le p(a|b), \tag{3.78}$$

for all probability measures $p:R \to [0,1]$ with $p(b)$, $p(d)>0$.

Also, define any $(a|b)$ to be a *contradiction for* CPL, written $\not\vdash_{CPL}(a|b)$, iff for all probability measures $p:R \to [0,1]$, with $p(b)>0$,

$$p(a|b) = 0. \tag{3.79}$$

It readily follows that

$$\not\vdash_{CPL}(a|b)\ \text{iff}\ a\&b = 0\ \text{iff}\ \vdash_{CPL}(a|b)'. \tag{3.80}$$

With a bit more difficulty, one can show (3.81)

$(c|d) \to_{CPL} (a|b)$ iff$(\vdash_{CPL}(a|b)$ or $\not\vdash_{CPL}(c|d)$ or $(c|d)\le(a|b))$.

(XI) Call $(a|b)$ and $(c|d)$ *semantically* (or *uniformly in probability*) *equivalent*, written $(c|d) \leftrightarrow_{CPL} (a|b)$, iff

$$p(a|b) = p(c|d), \tag{3.82}$$

for all probability measures $p:R \to [0,1]$ with $p(b)$, $p(d)>0$.

It follows that

$(c|d) \leftrightarrow_{CPL} (a|b)$ iff $((\vdash_{CPL}(a|b)$ and $\vdash_{CPL}(c|d))$ or

$(\not\vdash_{CPL}(a|b)$ and $\not\vdash_{CPL}(c|d))$ or

$(a|b) = (c|d))$. (3.83)

(XII) Finally, mention can be made of a weaker form of semantic implication by restricting probabilities in the above definitions to subclasses or even to a

single probability measure.

(iii) Another area of basic interest concerns higher order conditioning. Thus, if beginning with unconditional events say $a,b,c,d \in R$, it is meaningful to consider the conditional events $(a|b)$, $(c|d)$, then why should one not consider, in turn, the second order conditional event $((a|b)|(c|d))$? (This issue is addressed in [7], section 5 in some detail.) Basically, the appropriate definition for the above expression is the formal higher order analogue of (3.11). Thus

$$((a|b)|(c|d)) = \{(x|y):(x|y)\in\tilde{R},(x|y)\&(c|d)=(a|b)\&(c|d)\} \quad (3.84)$$

One major result ([7], Theorem 5.5) concerns the class union mapping $u:P(P(R)) \to P(R)$, where for any $\mathcal{Q} \subseteq P(R)$,

$$u(\mathcal{Q}) \overset{d}{=} \{x: x\in A\in\mathcal{Q} \text{, for any } A\in\mathcal{Q}\}. \quad (3.85)$$

Noting that u is a surjective homomorphism relative to all natural extensions of operations from $P(R)$ to $P(P(R))$, not just boolean operations, and noting that all boolean operations over $\tilde{R}$ ($\subseteq P(R)$) can be extended to $P(\tilde{R})$ ($\subseteq P(P(R))$) in the natural sense, though unlike the first order case, $\tilde{R}$, closure problems arise for the higher order case, the following reduction holds:

$$u(((a|b)|(c|d))) = (a|b\&((a'\&d')\vee(c\&d)) \text{ ,} \quad (3.86)$$

for all $a,b,c,d \in R$. In particular,

$$u(((a|b)|c) = u(((a|b)|(c|b))) = (a|b\&c). \quad (3.87)$$

Thus, through the above relations, in a sense, one need never consider higher order conditioning.

(iv) One can establish optimal approximations of arbitrary subclasses of R through conditional events, e., if $A \subseteq R$, the *best upper approximation of* A *through* $\tilde{R}$ satisfies the relation ([7], section 6)

$$A \subseteq \text{cond}(A) \overset{d}{=} \cap \{(a|b): A \subseteq (a|b) \in \tilde{R}\} \subseteq R, \quad (3.88)$$

noting trivially that for any $(a\ b) \in \tilde{R}$,

$$(a|b) = \text{cond}((a|b)). \quad (3.89)$$

The chief results include the following (see [7], Theorems 6.1, 6.2, and Corollary 6.3):

(I) If now R is a complete boolean algebra, then for any $A \subseteq R$,

$$\text{cond}(A) = (\&(A)|\&(A) \vee (v(A))') \ (\in \tilde{R}), \quad (3.90)$$

where

$$\&(A) \overset{d}{=} \underset{a\in A}{\&a} \text{ , } v(A) \overset{d}{=} \underset{a\in A}{va} \quad (3.91)$$

are the (possibly non-finite) extended definitions of & and v due to the completeness of R.

(II) Let $R = \mathcal{B}$, the class of all borel subsets of $\mathbb{R}$, the real (or one-dimensional euclidean) line. Let $g:\mathbb{R}^m \to \mathbb{R}$ be any function. Denote the natural extension of g to $P(\mathbb{R})$ restricted to $\mathcal{B}$ ($\subseteq P(\mathbb{R})$) as simply $g:\mathcal{B}^m \to \mathcal{B}$ (assuming g is sufficiently measurable). In turn, denote the natural extension to $P(\mathcal{B})$ restricted to $\tilde{\mathcal{B}}$, the conditional extension of $\mathcal{B}$, as simply $g:\tilde{\mathcal{B}}^m \to P(\mathcal{B})$. Then for all (borel sets) $a_i, b_i \in \mathcal{B}$, with $a_i \subseteq b_i$, $i=1,,,m$, noting as ordinary sets, $\subseteq$ replaces $\leq$, $\cap$ replaces &, $\cup$ replaces v, etc.,

$$\text{cond}(g((a_1|b_1),\ldots,(a_m|b_m))) = (\alpha|\beta), \quad (3.92)$$

where

$$\alpha \overset{d}{=} g(a_1,\ldots,a_m), \quad (3.93)$$

$$\beta \overset{d}{=} \alpha \cup (g(b_1\to a_1,\ldots,b_m \to a_m))' \text{ ,} \quad (3.94)$$

where here $b_i \to a_i = b_i' \cup a_i$.

If g is commutative and associative, then so is $\text{cond}(g):\tilde{\mathcal{B}}^m \to \tilde{\mathcal{B}}$.

The above development is particularly useful in determining the optimal approximations for naturally extended arithmetic operations, since these, unlike the boolean operations, in general do not possess closure properties over $\tilde{\mathcal{B}}$. These results lead, in sequence, to the development of random conditional events and related ideas. (See, in particular, [7], sections 7.3 and 10.4.)

4. PROBABILITY EVALUATIONS AND APPLICATIONS—

Following the brief overview of the role of data fusion in developing a generic model for typical C^3 processes in section 1 and the motivations and mathematical and computational structures for conditional events, this section presents the fundamental links for probability evaluation and application to combining of evidence and data fusion.

First note the following theorem :

Theorem 4.1. ([7], Theorems 1.5, 7.1, Remark 1.3)

Let $p:R \to [0,1]$ be a given probability measure. Then:

(i) With slightly additional conditions placed upon p (non-atomicity), the only possible extension of p to first $p:R_b \to [0,1]$, for each $b\in R$, and then to $p:\tilde{R} \to [0,1]$, where $p((a|b))$ is some fixed function (not dependent upon any given $a,b \in R$) of $p(a\&b)$ and $p(b)$ such that $p((\cdot|b)):R_b \to [0,1]$ is a probability measure is

$$p((a|b)) = p(a|b) = p(a\&b)/p(b), \quad (4.1)$$

i.e., ordinary conditional probability must be assigned to conditional events.

(ii) The extension $p:\tilde{R} \to [0,1]$ is *monotone increasing*, i.e., for any $a,b,c,d \in R$,

$$(a|b) \leq (c|d) \text{ implies } p(a|b) \leq p(c|d). \quad (4.2)$$

By using the demorgan property among others (see Theorem 3.7(iii) and Corollary 3.1), all computations for probabilities of (compound) boolean functions of conditional events can be reduced to computing probabilities of only conjunctions of conditional events. With this in mind, the following result shows that the measure-free conditional event approach presented here in conjunction with probability evaluations can, in a sense, be identified with a modified form of Approach B given in section 2.3, *with the joint probability measure* p_o, *in effect, determined through conditional event conjunction and initial probability measure* p *over* R:

Theorem 4.2 ([7], section 10.4, issue (x))

Let $p:R \to [0,1]$ be a given probability measure, let $b,d \in R$ be arbitrary with $p(b),p(d)>0$. For each $s,t \in \mathbb{R}$, consider the infinite left rays

$$a_s \overset{d}{=} (-\infty,s] \text{ , } c_t \overset{d}{=} (-\infty,t] \text{ .} \quad (4.3)$$

Then, $p((a_s|b)\&(c_t|d))$, as a function of (s,t), is a legitimate cumulative probability distribution function over $\mathbb{R}^2$ with $p((a_s|b)\&(d|d))$, as a function of s, and $p((c_t|d)\&(b|b))$, as a function of t, being marginal cumulative probability distribution functions over $\mathbb{R}$. ∎

BASIC PRINCIPLES FOR EVALUATING PROBABILITIES OF CONDITIONED INFORMATION

(1). Determine whether or not the evidence has truly differing antecedents.

(2). If the antecedents of the information are identical then apply the usual calculus of relations for PL. For example, suppose the same source, sensor or human, on two different occasions produces an estimate of target location and it is desired to obtain the probability of the disjunction, since if the resulting probability is sufficiently low, no further investigation will be

carried out. If the source error, though possibly independent, is relatively small, then both target estimates can be considered unconditioned information. Or, perhaps, by sheer coincidence, the same error causing mechanism is present and behaves the same way in producing the two estimates, which themselves can differ greatly in terms of random behavior and probability characteristics. Or, further, where Approach B is valid: the actual joint probabilities of the source random mechanism are known, one can then reduce all computations to that involving only PL, because of the common joint antecedent.

(3) But, in general, as in the simplified examples in section 2.1, the antecedents of arriving conditioned information arise from widely varying sources, of which little or nothing is known concerning joint probability distributions, yet relatively much is known concerning the individual (consequent) event probabilities.

(4) Isolate all the relative unconditional events a,b,c,d,e,f,.. making up the forms of the individual conditional probabilities $p_1(a|b)$, $p_2(c|d)$, $p_3(e|f)$,..

(5) Temporarily ignoring the probabilities, determine just what compound boolean operation is desired. Perhaps one simply is seeking to obtain, in some way, the "joint" probability of $(a|b)$, $(c|d)$, $(e|f)$, i.e. in actuality, the probability of the conjunction $(a|b)\&(c|d)\&(e|f)$, or any other combination as in the examples previously mentioned. Carry out the measure-free computations, based upon the boolean function desired, in accordance with the calculus of conditional event operations as given in section 3.3. Note that any such operation reduces the collection of, possibly many, conditional events to a single conditional event, say $(\alpha|\beta)$, where now α,β are known boolean functions of the input events a,b,c,d,e,f...

(6) Approximate by a single probability measure p, say, in place of all the differently arising probabilities in step 4, so that, in a consistent sense,

$p(a\&b)=p_1(a\&b)$, $p(c\&d)=p_2(c\&d)$, $p(e\&f)=p_3(e\&f)$, ..

Obviously, if the very same event, say b is assigned two distinct probabilities from two sources, say $p_1(b)$ and $p_2(b)$, some concensus must be determined before a final common assignment- by perhaps use of least squares, maximum likelihood, or maximal entropy techniques.

In lieu of step (4), it is possible that the conditional probabilities are all given relative to some common joint probability $p_i=p$, i=1,2,..., in which case the computational task simplifies.

(7) Simply use eq.(4.1). An example of a very general formula encompassing the probability evaluation of an arbitrary boolean function of multiple conditional events (consider again the comments at the beginning of Corollary 3.1) is given in (3.32).

(8) A related issue to the actual carrying out of the above steps in considering data fusion problems is that of bounding uncertainty or information. One straightforward result is the following, utilizing the results of section 3.3 and the well known Fréchet bounds for probabilities (see, e.g. [12]):

Theorem 4.3.

For any $a_j, b_j \in R$, j=1,...,n, and probability measure $p: R \to [0,1]$,

$$\max\left(0, \sum_{j=1}^{n} p(a_j|\underset{j=1}{\overset{n}{\&}} b_j)-(n-1)\right)\le p(\underset{j=1}{\overset{n}{\&}}(a_j|b_j))\le \min_{j=1,\dots,n}(p(a_j|b_j)), \tag{4.4}$$

$$\max_{j=1,\dots,n}(p(a_j|b_j))\le p(\underset{j=1}{\overset{n}{\vee}}(a_j|b_j))\le \min\left(1, \sum_{j=1}^{n} p(a_j|\underset{j=1}{\overset{n}{\&}} b_j)\right). \tag{4.5}$$

∎

Making the usual definitions for *information uncertainty* or *entropy*, for any $a,b \in R$, $p(b)>0$,

$$H_p(a|b) \overset{d}{=} -\log(p(a|b)) , \tag{4.6}$$

it follows that Theorem 4.3 can be converted immediately to bounds on information uncertainty.

Ideally, what is sought is the conditional event calculus -as developed here- analogue of Hailperin's results [12] concerning the bounding of probability values (and hence, correspondingly, information uncertainty) for arbitrary boolean functions of unconditional events when probability bounds are known for the individual (unconditional) events. However, at this point, one must be content with the rather simple results given in Theorem 4.3, until further results are obtained in this area. These properties also tie-in with CPL and the weakened forms of implication mentioned in section 3.5(ii)(XII). More details of this will be presented in future work.

(9) Generalizing the situation described in steps (6) and (7), one can have events of interest a_i, b_i, c_i, d_i, ..., ϕ_i (null event), Ω_i (universal event) all belonging to boolean algebra of events R_i corresponding to common probability measure $p_i: R_i \to [0,1]$, for i=1,2, say. But, by making the usual marginal identifications $a_1=a_1\times\Omega_2$, $b_1=b_1\times\Omega_2$, ..., $a_2=\Omega_1\times a_2$, $b_2=\Omega_1\times b_2$, ..., and assuming that the joint probability measure p_o of p_1 and p_2 is known, noting that because of (4.1),

$$p_1(a_1|b_1)=p_o(a_1\times\Omega_2|b_1\times\Omega_2),\dots, \tag{4.7}$$

$$p_2(a_2|b_2)=p_o(\Omega_1\times a_2|b_2\times\Omega),\dots , \tag{4.8}$$

it follows that all of the previous steps are now valid for the situation here with p_o replacing p, and the marginal identifications for the events.

(10) Higher order conditioning, i.e., when the events in the above steps such as a,b,c,... are in actuality in conditioned form already (which may well be a common situation), can be treated in a straightforward way by use of the results in section 3.5(iii).

(11). Finally, it should be remarked that for non-stochastic information, such as that containing linguistic-based evidence, as considered, e.g., in [11], an analogous calculus of conditional forms can be developed, based upon (3.11) and the natural extensions of operations [13]. The corresponding full ALDP's should prove of use in treating combination of evidence problems and data fusion in general.

5. ACKNOWLEDGMENTS

This work was supported jointly by the Naval Ocean Systems Center Program for Independent Research (IR) and the Joint Directors Of Laboratories, Technical Panel for C^3, Basic Research Group (JDL,TPC^3,BRG).

6. REFERENCES

1. Goodman, I.R., "A general theory for the fusion of data", *Proc. 1st Tri-Serv. Data Fus. Symp.* 1987, 254-270.
2. Goodman, I.R., "A probabilistic/possibilistic approach to modeling C^3 systems: Part II", *Proc. 1st Symp. C^3 Research*, Dec., 1987, 41-48.
3. Goodman, I.R., "A probabilistic/possibilistic approach to modeling C^3 systems", *Proc. 9th MIT/ONR Wrkshp. C^3 Sys.*, Dec., 1986, 53-58.
4. Goodman, I.R. & Nguyen, H.T., *Uncertainty Models for Knowledge-Based Systems*, North-Holland, Amst., 1985.
5. Goodman, I.R., "Toward a general theory of C^3 processes", *Proc. 2nd Symp. C^3 Research*, to be publ., 1988.
6. Goodman, I.R., "A measure-free approach to conditioning", *Proc. 3rd AAAI Wrksp. Uncert. in AI*, 1987, 270-277.
7. Goodman, I.R. & Nguyen, H.T., *A Theory of Measure-Free Conditioning*, submitted for publ. (monograph) 1988-1989 .
8. Calabrese, P., "An algebraic synthesis of the foundations of logic and probability", *Inf. Sci.* 42, '87, 187-237.
9. Schay, G., "An algebra of conditional events", *JMAA* 24(2), '68, 334-344.
10. Copeland, A.H., "Implicative boolean algebra", *Math. Z.* 53(3), '50, 285-290.
11. Goodman, I.R., *PACT*, NOSC Tech Doc. 878, March, 1986.
12. Hailperin, T., "Probability Logic", *Notre Dame J. Formal Logic*, 25(3), 1984, 198-212.
13. Goodman, I.R., "A unified approach to the modeling of uncertainties", to be submitted, 1988.